Reply to: "Reliability and Validity of Speech Evaluation in Adductor Spasmodic Dysphonia; Common Mistake and Statistical Issues" by Saori Yanagida

To the Editor,

We thank Dr. Sabour for his interest in our article "Reliability and validity of speech evaluation in adductor spasmodic dysphonia."[1] In this research, Pearson correlation coefficient was obtained for the purpose of evaluating the intra- and interrater or intra- and intermeasurer reliabilities for perceptual evaluation and acoustic measures. As Dr. Sabour pointed out, calculation of the intraclass correlation coefficient (ICC) showed there was almost no difference between that value and Pearson correlation coefficient (Tables 1, 2). There are several ways to evaluate the reproducibility of quantitative variables[2–4]; therefore, we do not think that it is necessary to calculate ICC in our paper. Regarding the validity assessment, we used perceptual evaluation and acoustic measures in this study to evaluate whether it is useful as an evaluation method to complement the diagnosis of spasmodic dysphonia. Calculation of the sensitivity and specificity to examine the validity of an evaluation scale is a method commonly used in previous studies,[5,6] and we think that the statistical analysis used in this study is appropriate. We believe that our method and result do not lead to the misdiagnosis and mismanagement of patients pointed out by Dr. Sabour. Thank you for affording us the opportunity to respond.

Saori Yanagida
Department of Communication Disorders, School of Rehabilitation Sciences, Health Sciences University of Hokkaido, Hokkaido, Japan

https://doi.org/10.1016/j.jvoice.2017.10.008

REFERENCES

1. Yanagida S, Nishizawa N, Hashimoto R, et al. Reliability and validity of speech evaluation in adductor spasmodic dysphonia. *J Voice*. 2017;doi:10.1016/j.jvoice.2017.06.022. in press.
2. Rabinov CR, Kreiman J, Gerratt BR, et al. Comparing reliability of perceptual ratings of roughness and acoustic measures of jitter. *J Speech Hear Res* 1995;38:26–32.
3. Eadie TL, Baylor CR. The effect of perceptual training on inexperienced listeners' judgments of dysphonic voice. *J Voice*. 2006;20:527–544.
4. Belafsky PC, Postma GN, Koufman JA. The validity and reliability of the reflux finding score (RFS). *Laryngoscope*. 2001;111:1313–1317.
5. Bjelland I, Dahl AA, Haug TT, et al. The validity of the hospital anxiety and depression scale: an updated literature review. *J Psychosom Res*. 2002;52:69–77.
6. Barber PA, Demchuk AM, Zhang J, et al. Validity and reliability of a quantitative computed tomography score in predicting outcome of hyperacute stroke before thrombolytic therapy. *Lancet* 2000;355:1670–1674.

TABLE 1.
Intra- and Interrater Reliability for Perceptual Evaluation Using ICC

	Overall Severity	Strangulation	Interruption	Tremor	Strained Speech
Intra-rater					
Rater1	.941	.803	.918	.886	.868
Rater2	.978	.934	.954	.909	.910
Rater3	.958	.974	.903	.941	.913
Inter-rater					
First session	.913	.874	.850	.770	.850
Second session	.916	.873	.859	.831	.839

TABLE 2.
Intra- and Intermeasurer Reliability for Acoustic Events Using ICC

	Percentage (%)		
	Frequency Shifts	Aperiodic Segments	Phonation Breaks
Intrameasurer	.416	.958	.870
Intermeasurer (average)	.718	.740	.902

0892-1997